
CAPSTONE PROJECT

CUSTOMER SEGMENTATION USING MACHINE LEARNING

Presented By :

- **Anurag Gupta**
- **Kanpur Institute of Technology**
- **Artificial Intelligence and Machine Learning**

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Results
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

- Customer segmentation is crucial for businesses to tailor their marketing strategies and improve customer satisfaction.
- Aim: Segment customers based on tenure, monthly charges, and total charges using machine learning techniques.

PROPOSED SOLUTION

❑ Components:

- Data Collection : Gathered data from 'Customer_Segmentation.csv' containing tenure, monthly charges, total charges, and churn status.
- Data Preprocessing : Handled missing values and outliers, filled missing values, and standardized numerical features.
- Exploratory Data Analysis (EDA) ; We visualized the relationship between tenure and monthly charges using a scatter plot and created a pie chart to show the distribution of tenure for a subset of customers.
- Feature Engineering : We standardized the numerical features (tenure, monthly charges, total charges) using Standard Scaler and normalized the numerical features using Min Max Scaler.
- Model Building : We split the standardized features and the target variable ('Churn') into training and test sets. Built a Random Forest classifier model using the standardized features and trained the model on the training set and evaluated its performance on the test set.

SYSTEM DEVELOPMENT APPROACH

- Exploratory Data Analysis (EDA):
 - Scatter plot for tenure vs. monthly charges.
 - Pie chart for tenure distribution.
 - Line graph for tenure vs. total charges.
- Feature Engineering:
 - Standardized and normalized numerical features using StandardScaler and MinMaxScaler.

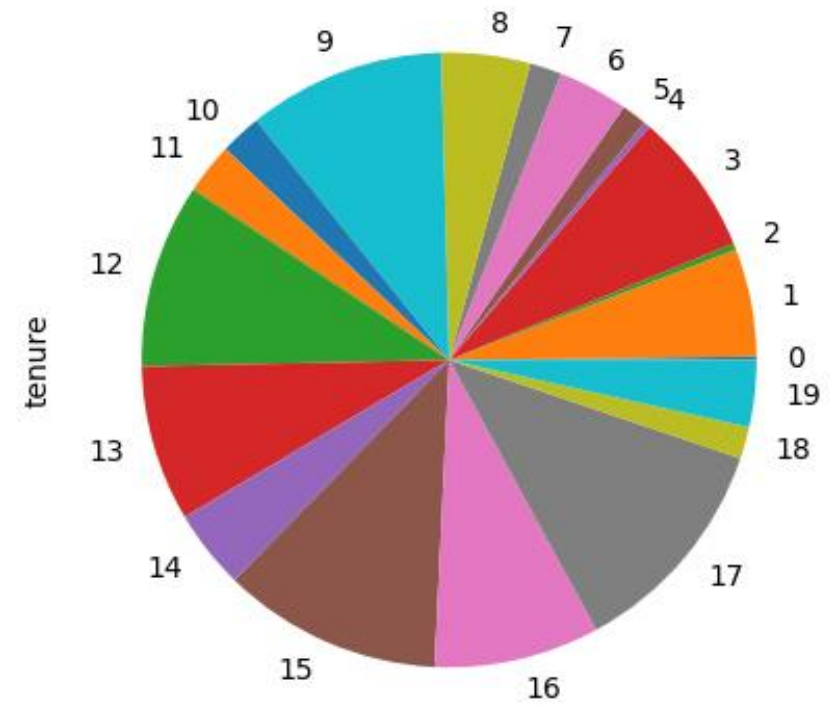
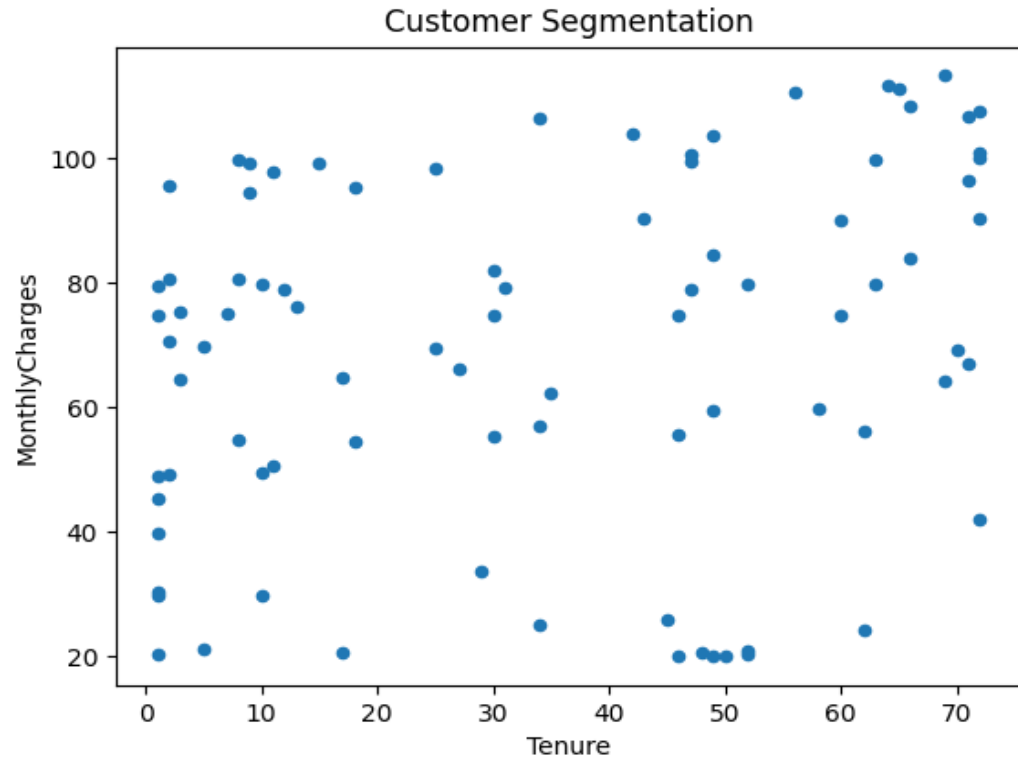
ALGORITHM & DEPLOYMENT

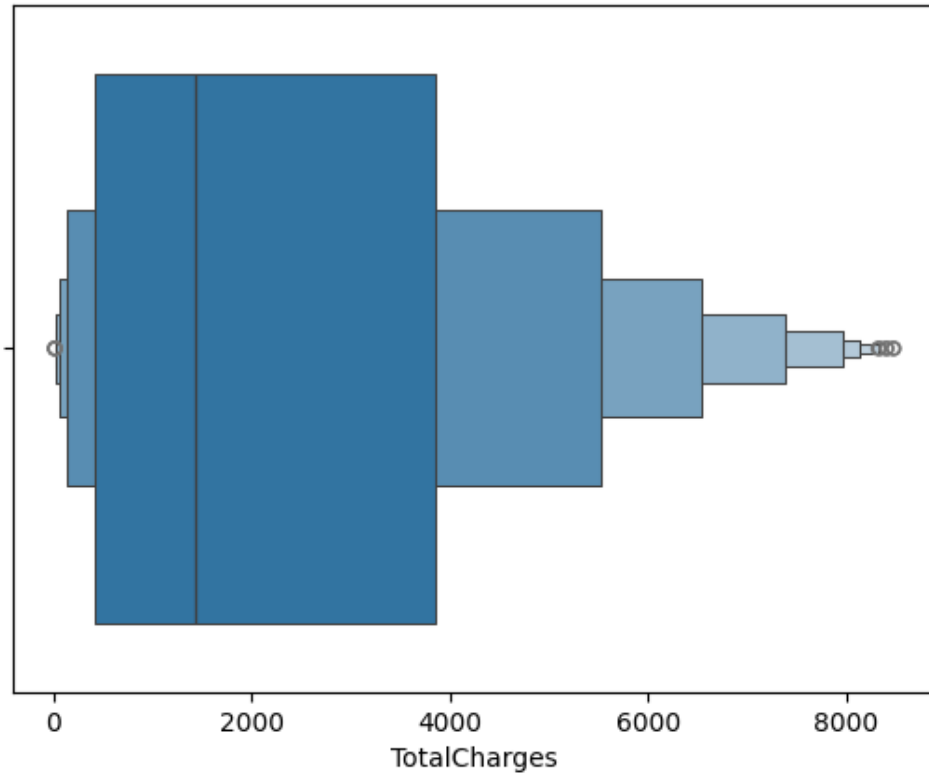
□ Python libraries used in different fields

- Model Building : sklearn.model_selection
- Split data into training and test sets : sklearn - train_test_split
- Built a Random Forest classifier using standardized features : sklearn.ensemble
- Evaluated the model : sklearn.metrics
- Data Visualizations : Matplotlib.pyplot
- Addition required : Seaborn, Pandas, Numpy, LogisticRegression, StandardScaler, MinMaxScaler

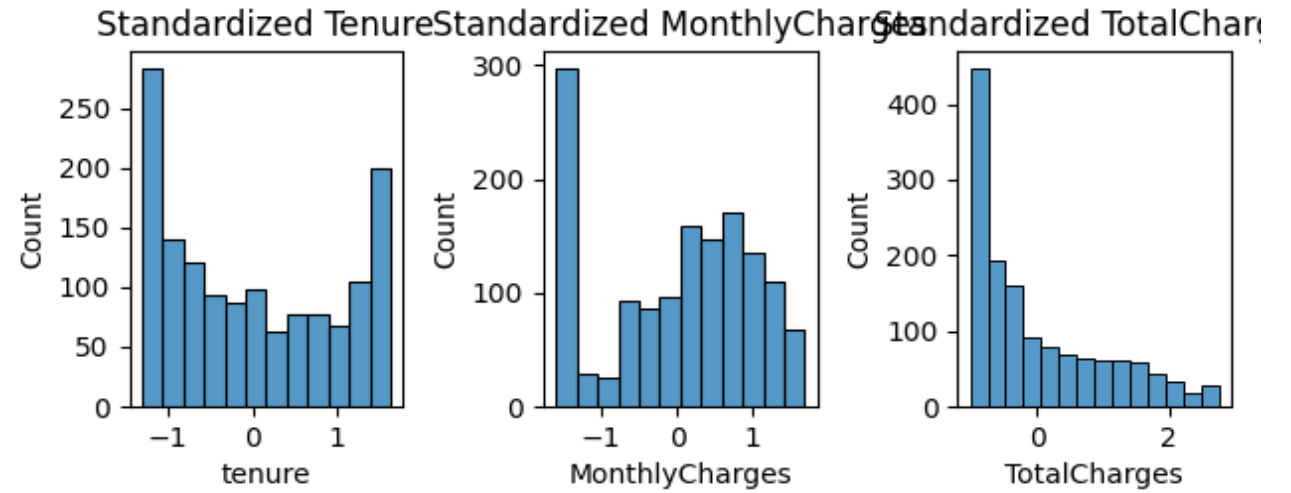
RESULTS

- Model Performance:
- Achieved an accuracy of [77.74%].
- Acquire the prediction Model.
- Insights from model evaluation and performance metrics.
- Data Visualization in next slides.





Describe the Outliers



CONCLUSION

- Summary :
- Importance of customer segmentation for businesses.
- Machine learning aids in automating segmentation and improving marketing strategies.
- Future Work:
- Further analysis to improve model performance.
- Exploration of other segmentation techniques.

FUTURE SCOPE

- Potential enhancements:
- Incorporating additional data sources.
- Optimizing algorithms for better performance.
- Expanding the system to cover multiple regions.

REFERENCES

- Relevant sources, research papers, and articles used for developing the solution.
- Source Code - [Link](#)
- CSV File – [Link](#)
- Testimonial Template - [Link](#)

COURSE CERTIFICATE 1

In recognition of the commitment to achieve
professional excellence



Anurag Gupta

Has successfully satisfied the requirements for:

Getting Started with Enterprise-grade AI



Issued on: 09 JUL 2024

Issued by IBM

Verify: <https://www.credly.com/go/gVa7yqa0>



COURSE CERTIFICATE 2

In recognition of the commitment to achieve
professional excellence



Anurag Gupta

Has successfully satisfied the requirements for:

Machine Learning for Data Science Projects



Issued on: 10 JUL 2024

Issued by IBM

Verify: <https://www.credly.com/go/t8Y3L1P9>





THANK YOU